

⑩ 日本国特許庁(J.P.)

⑪ 特許出願公開

⑫ 公開特許公報(A) 平1-159891

⑬ Int.Cl.⁴

G 11 C 11/34

識別記号

3 6 2

庁内整理番号

C-8522-5B

⑭ 公開 平成1年(1989)6月22日

審査請求 未請求 発明の数 1 (全10頁)

⑮ 発明の名称 半導体記憶装置

⑯ 特 願 昭62-322126

⑰ 出 願 昭62(1987)12月17日

⑱ 発 明 者 藤 島 一 康 兵庫県伊丹市瑞原4丁目1番地 三菱電機株式会社エル・エス・アイ研究所内

⑲ 発 明 者 松 田 吉 雄 兵庫県伊丹市瑞原4丁目1番地 三菱電機株式会社エル・エス・アイ研究所内

⑳ 発 明 者 朝 倉 幹 雄 兵庫県伊丹市瑞原4丁目1番地 三菱電機株式会社エル・エス・アイ研究所内

㉑ 出 願 人 三菱電機株式会社 東京都千代田区丸の内2丁目2番3号

㉒ 代 理 人 弁理士 大岩 増雄 外2名

明 開 書

1. 発明の名称

半導体記憶装置

2. 特許請求の範囲

(1) 行および列状に配列されて各々が情報を記憶するメモリセルからなるメモリセルアレイに対し、行アドレスおよび列アドレスを指定することで前記メモリセルより情報を取り出す半導体記憶装置であって、

行アドレス指定された1行分のメモリセルの情報を検知し格納するセンスアンプと、

このセンスアンプ内の各1ビット情報に対しそれぞれ複数個のメモリセルを有し、前記センスアンプ内の各情報を選択的にメモリセルに取り込むスタティックメモリセルアレイと、

前記スタティックメモリセルアレイあるいは前記メモリセルアレイのいずれにアクセスするかを選択するスイッチ手段と、

前記スタティックメモリセルアレイのいずれのメモリセルにアクセスするかを選択するウェイデ

コードとを備えた半導体記憶装置。

(2) 前記スタティックメモリセルアレイは前記センスアンプ内の情報を予め定められたブロック単位で取り込み、前記スタティックメモリセルアレイにおけるいずれのブロックにアクセスするかを選択するブロックデコードをさらに備えた特許請求の範囲第1項記載の半導体記憶装置。

(3) 前記ウェイデコードは前記スタティックメモリセルアレイにおける全てのメモリセルを非活性にする機能をさらに備えた特許請求の範囲第1項または第2項記載の半導体記憶装置。

3. 発明の詳細な説明

(産業上の利用分野)

この発明はキャッシュメモリを内部に有する半導体記憶装置に関する。

(従来の技術)

従来、コンピュータシステムのコストパフォーマンスを向上させるため、低速だが低コストで大容量なダイナミックRAM(DRAM)をメインメモリに使用し、このメインメモリとCPU間に

高速なバッファとして、小容量の高速メモリを設けることが、よく行われていた。上記した高速バッファはキャッシュメモリと呼ばれ、CPUが必要とするようなデータのブロックをメインメモリからコピーし、保持している。CPUがアクセスするアドレスのデータがキャッシュメモリ内に存在する時(キャッシュヒット)、CPUは必要とするデータをキャッシュメモリより取り込む。一方、CPUがアクセスするアドレスのデータがキャッシュメモリ内に存在しない時(キャッシュミス)、CPUは低速なメインメモリ(DRAM)より、必要とするデータを取り込む。

上記したキャッシュメモリシステムをメモリシステムに組み込むには、高価な高速メモリを必要とするのでコストを重視する小型のコンピュータシステムでは使用することができなかった。そこで、DRAMの有しているページモード、スタティックコラムモード等の高速アクセス機能を利用し、簡易なキャッシュシステムを構成していた。

以下、第5図の波形図を参照して、ページモ

ド、スタティックコラムモードの説明を行う。同図において(a)は通常のDRAMのサイクル、(b)はページモードサイクル、(c)はスタティックコラムモードサイクルである。

同図(a)に示すように、通常サイクルでは、信号RAS (Row Address Strobe)の降下エッジでマルチプレクスアドレス信号MAより行アドレス(Row Address) RAをDRAM内に取込み、信号CAS (Column Address Strobe)の降下エッジでマルチプレクスアドレス信号MAより列アドレス(Column Address) CAをDRAM内に取り込む。そして、行アドレスRA、列アドレスCAにより選択されたメモリセルのデータをデータ出力 D_{out} として得る。通常サイクルは上記したサイクルでデータを読み出すため、アクセス時間としては信号RASの降下エッジ時からデータ出力 D_{out} が有効になるまでの時間 t_{RAC} (RASアクセスタイム)を要する。このアクセス時間 t_{RAC} は、通常100ns程度である。なお、 t_{RP} は信号RASのプリチャージ時間、 t_C はサ

イクル時間であり、通常 $t_C = 200ns$ 程度である。

同図(b)に示すように、ページモードサイクルでは同一行アドレスRA上で複数の列アドレスCAでデータの読出しが行える。従って、アクセス時間は信号CASの降下エッジ時からデータ出力 D_{out} が有効になるまでの時間 t_{CAC} (CASアクセスタイム)となり、通常サイクルでのアクセス時間 t_{RAC} の半分程度の時間となり、通常50ns程度である。なお、 t_{CP} は信号CASのプリチャージ時間、 t_{PC} はサイクル時間である。

同図(c)に示すように、スタティックコラムモードではページモードの信号CASの立下りエッジを不要にし、列アドレスCAをあたかもスタティックRAMのように動作させている。従ってアクセス時間はマルチプレクスアドレスMA変化時からデータ出力 D_{out} が有効になるまでの時間 t_{AA} (アドレスアクセスタイム)となり、 t_{CAC} 同様通常サイクルでのアクセス時間 t_{RAC} の半分程度となり、通常50ns程度である。

第6図は、ページモードあるいはスタティックコラムモードが可能な従来のDRAM素子の基本構成を示す構成ブロック図である。

同図に示すように、行アドレスバッファ1、列アドレスバッファ2がマルチプレクスアドレス信号MAより各々行アドレスRA、列アドレスCAを取込んでいる。そして信号RASの降下エッジが行アドレスバッファ1に入力されると、行アドレスRAが行デコーダ3へ送られ、次段のワードドライバ4を駆動することで、行アドレスRAにより選択されたメモリセルアレイ5内の1本のワード線(図示せず)を活性化する。

そして、活性化されたワード線に接続された全メモリセルのデータが、メモリセルアレイ5内の全ビット線(図示せず)を介してセンスアンプ6へ送られる。センスアンプ6は得られたデータを検知し、増幅する。したがって、この時点で指定された行アドレスRA一行分のデータがセンスアンプ6にラッチされている。以降、行アドレスRAが同一のデータをアクセスする場合は、前述し

たページモード、スタティックコラムモードが利用できる。

つまり、ページモードでは、信号CASの両下エッジが列アドレスバッファ2に入力されると、列アドレスCAが列デコーダ7に送られ、センスアンプ6に格納されているデータ群のいずれかを有効にすることで、出力バッファ8を介してデータ出力D_{out}を得る。スタティックコラムモードの場合も起動をマルチプレクスアドレスMAの変化による点を除き同様の動作を行う。なお、9はデータの入出力を制御するI/Oスイッチ、10は入力バッファ、D_{in}はデータ入力である。

第7図はページモード（あるいはスタティックコラムモード）を利用した簡易キャッシュシステムを有する従来のメモリシステムのブロック構成図である。同図に示すように、このメモリシステムは8個の1M×1構成のDRAM素子11～18を使用し構成した1Mバイトのメモリシステムである。従ってアドレス線は20本（ $2^{20}=1048576=1M$ ）必要とするが、実際上はアド

レスマルチプレクサ21より行アドレスRA（10ビット）、列アドレスCA（10ビット）に分けたマルチプレクスアドレス信号MAが送られる10本のアドレス線が各々のDRAM素子11～18に接続されている。

第8図は、第7図で示したメモリシステムのキャッシュ動作を示した波形図である。以下、第8図および第6図を参照しつつ第7図のメモリシステムの動作を説明する。なお、ラッチ22には、既に直前にアクセスされた行アドレスRA1がラッチされており、センスアンプ6内には行アドレスRA1の全データが既にラッチされているとする。

このような状態で、CPU26が必要とするデータの20ビットのアドレス信号A₀をアドレスジェネレータ23より発生する。このアドレス信号A₀から行アドレスRA2がコンパレータ24に入力され、コンパレータ24はこの行アドレスRA2とラッチ22に格納されている行アドレスRA1との比較を行い、RA1=RA2であれば、

センスアンプ6に保持しているデータ群にアクセスされた（キャッシュヒット）ことになり、コンパレータ24は活性化した（“H”レベル）キャッシュヒット信号CH（Cache Hit）をステートマシン25に送る。活性化した信号CHを受けたステートマシン25は信号RASを“L”レベルに保ったまま、信号CASをトグルする（立ち上げた後に立ち下げる）ページモード制御を行い、アドレスマルチプレクサ21はDRAM素子11～18にマルチプレクスアドレスMAとして、列アドレスCAを供給し、各DRAM素子11～18のセンスアンプ6に格納されたデータ群より、列デコーダ7により選択されたデータを取り出す。このようにキャッシュヒットした場合、DRAM素子11～18から高速なアクセス時間t_{CAC}で、出力データD_{out}が得られる。

一方、コンパレータ24において、RA1≠RA2が判定されると、センスアンプ6に保持しているデータ群以外にアクセスされた（キャッシュミス）ことになり、コンパレータ24はステート

マシン25に非活性（“L”レベル）の信号CHを発生する。この時、ステートマシン25は信号RAS、CASの順にトグルする通常サイクルのDRAM素子11～18の制御を行い、アドレスマルチプレクサ21は行アドレスRA2、列アドレスCAの順にマルチプレクスアドレスMAをDRAM素子11～18に供給する。このようにキャッシュミスした場合、信号RASを第8図に示すようにプリチャージし、さらにDRAM素子11～18から低速なアクセス時間t_{RAC}で出力データD_{out}が得られることになる。このため、ステートマシン25はウェイト信号Waitを発生し、CPU26に待機をかける。また、ラッチ22はコンパレータ24より活性化されないキャッシュヒット信号CHを受けると新しい行アドレスRA2を保持する。

（発明が解決しようとする問題点）

従来の簡易キャッシュシステムは以上のようにセンスアンプ6によりラッチする形式で構成されているので、エントリー数は1である。従って、

同じ行アドレスRAに連続してアクセスする場合のみにキャッシュヒットとなるため、例えば連続する2つの行アドレスにまたがったプログラムルーチンが繰り返し実行される場合などには、必ずキャッシュミスが生じてしまうことになり、キャッシュヒット率が低いという問題点があった。

この発明は、上記した問題点を解決するためになされたもので、キャッシュヒット率を向上させた簡易キャッシュシステムを有する半導体記憶装置を得ることを目的とする。

(問題点を解決するための手段)

この発明に係る半導体記憶装置は行および列状に配列されて各々が情報を記憶するメモリセルからなるメモリセルアレイに対し、行アドレスおよび列アドレスを指定することで前記メモリセルより情報を取り出す方式であって、行アドレス指定された1行分のメモリセルの情報を検知し格納するセンスアンプと、このセンスアンプ内の各1ビット情報に対しそれぞれ複数個のメモリセルを有し、前記センスアンプ内の各情報を選択的にメモ

リセルに取り込むスタティックメモリセルアレイと、前記スタティックメモリセルアレイあるいは前記メモリセルアレイのいずれにアクセスするかを選択するスイッチ手段と、前記スタティックメモリセルアレイのいずれのメモリセルにアクセスするかを選択するウエイデコーダとを備えて構成されている。

(作用)

この発明におけるスタティックメモリセルアレイはセンスアンプ内の1ビット情報に対し複数個のメモリセルが設けられているため、異なる行アドレス上のデータを保持することができる。

(実施例)

第1図はこの発明の一実施例であるキャッシュ機能を有するメモリシステムのDRAM素子の基本構成を示すブロック構成図である。図面において1~4、8~10及びCAS、RAS、MA、RA、CA、CHは従来と同じであるので説明は省略し、以下従来と異なる点について述べる。

図面に示すようにメモリセルアレイ5をブロッ

クB1~B4と4分割して使用するため、センスアンプ6、1/0スイッチ9間にブロックB1~B4に対応してトランスファゲート31(31a~31d)、スタティックメモリセルアレイ32(32a~32d)を挿入している。トランスファゲート31は、第2図の詳細ブロック構成図に示すようにブロックデコーダ34により各々が制御されるため、その導通・非導通により、メモリセルアレイ5のデータをブロック(B1~B4)単位で、センスアンプ6を介して対応のスタティックメモリセルアレイ32a~32dへ転送が可能となる。

スタティックメモリセルアレイ32は、第2図に示すように、センスアンプ6に格納された1ビット情報に対し、トランスファゲート31を介して4個のスタティックメモリセル32_{a1}~32_{a4}を設けている。これらのメモリセル32_{a1}~32_{a4}は、ウエイデコーダ35の出力線W₁~W₄が“H”レベルの時活性化する。ウエイデコーダ35は第1図で示すようにウエ

イアドレスバッファ36を介して入力されるウエイアドレスWAをデコードし、出力線W₁~W₄を選択的に“H”レベルに立上げる。

ブロックデコーダ34a~34dは、各々列アドレスCAの上位2ビットと信号CHの反転信号を入力信号とするアンドゲートG1によりその活性化が制御される。つまり、信号CHが“L”レベルで、列アドレスCAの上位2ビットで選択されたブロックデコーダ34a~34dのいずれかが活性化し、信号CHが“H”レベルでは、どのブロックデコーダ34a~34dも活性化しない。またブロックデコーダ34a~34dのいずれかが活性化すると対応するトランスファゲート31a~31dが導通する。一方、列デコーダ7は列アドレスCAを入力信号とし、1/0スイッチ9のいずれか1つを有効にする。

第3図はこの発明の一実施例であるキャッシュ機能を有するメモリシステムを示したブロック構成図である。図面に示すように、従来と異なり、4ブロック、4スタティックメモリセル32_{a1}~

32_{B2}構成であるため、16個のラッチ22a~22p。(22a~22dはブロックB1、22e~22hはブロックB2、22i~22lはブロックB3、22m~22pブロックB4)を設けている。また、これらのラッチ22a~22pは各々行アドレスRAとスタティックメモリセル32_{B1}~32_{B4}のいずれを選択すべきかを示すウェイアドレスWAを格納しており、アドレス信号A_dの列アドレスCAの上位2ビットによりブロック単位に4個選択される。

コンパレータ24は行アドレスを取り込み、ラッチ22a~22pの中から選択された4個の格納された行アドレスと比較し、1個のラッチと一致すればキャッシュヒットとみなし、活性化した("H"レベルの)キャッシュヒット信号CHを出力し、同時に一致したラッチ22に格納されたウェイアドレスC_{WA}をウェイロジック37に出力する。一方、選択された4個のラッチ全てと一致しなかった場合、キャッシュミスとみなし非活性("L"レベルの)キャッシュヒット信号CHを

出力する。

ウェイロジック37はコンパレータ24から出力されるキャッシュヒット信号CHとウェイアドレスC_{WA}を入力信号とし、キャッシュヒット時には、入力されたウェイアドレスC_{WA}をそのままウェイアドレスWAとして、各DRAM11~18に出力する。一方、キャッシュミス時には、所定のアルゴリズムに従い決定されたウェイアドレスWAを各DRAM11~18及びラッチ22a~22pに出力する。先に述べた所定のアルゴリズムとは、例えば単純な先入れ、先出し方式(FIFO: First-in, First-out)、あるいは最後にアクセスされた時刻が最も古いものを追い出す方式(LRU: least recently used)等が考えられる。

以下、第4図のキャッシュヒット、キャッシュミス時の波形状を参照しつつ、第1図~第3図で示したこの発明の一実施例であるメモリシステムの動作を説明する。なお、ラッチ22a~22pには、既に各ブロックB1~B4の各スタティックメモリセル32_{B1}~32_{B4}において直前にアク

セスされた行アドレスRA1a~RA1p及び各ラッチ22a~22pに該当するウェイアドレスが各々ラッチされており、スタティックメモリセルアレイ32a~32pの各メモリセル32_{B1}~32_{B4}にはその時のブロックB1~B4ごとの全データが既にラッチされているとする。

このような状態で、図示しないCPUが要求とする20ビットのアドレス信号A_dをアドレスジェネレータ23より発生する。このアドレス信号A_dから行アドレスRA2がコンパレータ24に入力される。一方、アドレス信号A_dの列アドレスCAの上位2ビットにより選択されたブロックB1~B4に該当するラッチ22a~22pのいずれか4個のみを有効にする。ここで、説明の都合上ブロックB2のラッチ22e~22hが選択されたとすると、コンパレータ24は入力された行アドレスRA2とラッチ22e~22hに格納されている行アドレスRA1e~RA1h各々との比較を行い、RA1e~RA1hのいずれかがRA2と一致すれば、キャッシュヒットとみなし、

活性化した("H"レベルの)キャッシュヒット信号CHをステートマシン25、ウェイロジック37及び各DRAM素子11~18に送り、RA2と一致した行アドレスを格納したラッチ22e~22hに格納されたウェイアドレスを信号C_{WA}としてウェイロジック37に送る。"H"レベルのキャッシュヒット信号CHを受けたウェイロジック37は、入力されたウェイアドレスC_{WA}をそのままウェイアドレスWAとして各DRAM11~18に出力する。

この時、信号CHは"H"レベルとなるため、全てのブロックデコーダ34は活性化せず、全トランスファゲート31は導通せず、スタティックメモリセルアレイ32とセンスアンプ6間は電気的に遮断されている。

一方、ステートマシン25は信号CASをトグルするページモード制御を行ない、アドレスマルチプレクサ21はDRAM素子11~18にマルチプレクスアドレスMAとして列アドレスCAを供給する。一方、各DRAM素子11~18に入

力されたウェイアドレスWAはウェイアドレスバッファ36を介しウェイデコーダ35に入力される。そして、ウェイデコーダ35により、ウェイアドレスWAをデコードすることで出力線W₁ (i=1~4のいずれか)が“H”レベルに立上る。すると、各DRAM素子11~18のスタティックメモリセルアレイ32b内の“H”レベルに立ち上げられた出力線W₁に接続されたメモリセル32_{a1}より列デコーダ7により選択された列アドレスのデータをI/Oスイッチ9を介して取り出す。このようにしてキャッシュヒットした場合、DRAM素子11~18から高速なアクセス時間t_{CAC}で出力データD_{out}が得られる。

また、コンパレータ24においてRA2とRA1e~RA1hのすべてが不一致と判定されると、キャッシュミスとみなし、非活性(“L”レベル)のキャッシュヒット信号CHをステートマシン25、ウェイロックス36及び各DRAM素子11~18に送る。“L”レベルのキャッシュヒット信号CHを受けたウェイロックス37は所定のアルゴリズムにより決定されたウェイアドレスWA

を各DRAM素子11~18及びラッチ22a~22dに出力する。

この時、信号CHは“L”レベルとなるため、ブロックデコーダ34bのみ活性化され、トランスファゲート31bは導通し、スタティックメモリセルアレイ32bとセンスアンプ6間は電気的に接続される。なお、他のスタティックメモリセルアレイ32a、32c、32dとセンスアンプ6間は電気的に遮断されたままである。

一方、ステートマシン25は、信号RASを立ち下げ次に信号CASを立上げるサイクルでDRAM素子11~18の制御を行い、アドレスマルチプレクサ21は行アドレスRA2、列アドレスCAの順にマルチプレクスアドレスMAをDRAM素子11~18に供給する。一方、ウェイアドレスWAはウェイアドレスバッファ36を介し各DRAM素子11~18のウェイデコーダ35に入力される。そして、ウェイデコーダ35により、ウェイアドレスWAをデコードすることで出力線

W₁ (i=1~4のいずれか)が“H”レベルに立上る。そして、メモリセルアレイ5よりセンスアンプ6、トランスファゲート31b及びスタティックメモリセルアレイ32b、I/Oスイッチ9及び出力バッファ8を介して、列デコーダ7より選択されたデータを出力データD_{out}として読み出す。同時に、スタティックメモリセルアレイ32b内の“H”レベルに立ち上げられた出力線W₁に接続されたメモリセル32_{a1}内のデータを再換える。このようにキャッシュミス時には、DRAM素子11~18から低速なアクセス時間t_{RAC}で出力データD_{out}が得られることになる。このため、ステートマシン25はウェイト信号Waitを発生し、CPU26に待機をかける。また、列アドレスCAの上位2ビットとウェイアドレスWAにより選択されたラッチ22e~22hのいずれかには、新しい行アドレスRA2が保持される。(他のラッチ22の値は変化しない。)

このように、1行分の情報を保持したセンスアンプ6内の1ビット情報を4個のスタティックメ

モリセル32_{a1}~32_{a4}のいずれかに格納することができスタティックメモリセルアレイ32を設けたため、エンタリー数は4である。その結果、連続する2つの行アドレスにまたがったプログラムルーチンが繰り返し実行される場合などにも対応することができ、キャッシュヒット率は向上する。

さらに、キャッシュヒット、キャッシュミス時におけるDRAM素子11~18のメモリ管理をブロックB1~B4単位で行えるようにしたため、各ブロックB1~B4各々が独立して行アドレスに対するデータ群をスタティックメモリセルアレイ32a~32dに格納することができるので、エンタリー数は4×4=16となり、なお一回キャッシュヒット率は向上する。

なお、第1図で示した実施例ではメモリの読み出し、書き込みに関係なく、例えばウェイデコーダ35に読み込み信号WEを入力し、読み込み時(WE=“L”)は、全出力線W₁~W₄を非活性(“L”レベル)に設定することで、信号CHの“H”

・L、ウェイアドレスWAの値にかかわらず、スタティックメモリセルアレイ32内の全メモリセル32_{a1}～32_{a4}を活性化しないようにすることもできる。

なお、この実施例では、センスアンプ6の1ビット情報に対し4個のスタティックメモリセル32_{a1}～32_{a4}のいずれかに格納する例を述べたが、この数は適当に増減することができる。また、この実施例ではメモリセルアレイ5を4ブロックB1～B4構成としたが、ブロックの分割数も適当に増減することは勿論可能である。

(発明の効果)

以上説明したように、この発明によればスタティックメモリセルアレイはセンスアンプ内の1ビット情報に対し複数個のメモリセルが設けられたため、異なる行アドレス上のデータを保持することができる。その結果、エントリーを増加することができるため、キャッシュヒット率を向上させることができる。

4. 図面の簡単な説明

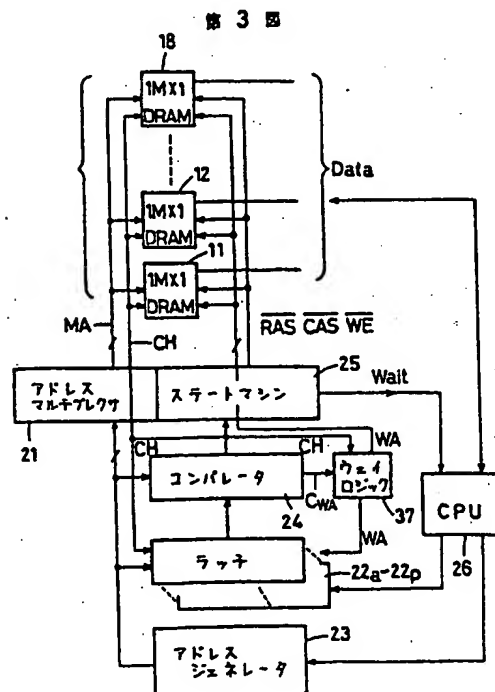
第1図はこの発明の一実施例であるキャッシュ機能を有するメモリシステムにおけるDRAM素子の構成説明図、第2図は第1図のDRAM素子の詳細な構成説明図、第3図はこの発明の一実施例であるキャッシュ機能を有するメモリシステムのブロック構成図、第4図はこの発明の一実施例のキャッシュ動作を示す波形図、第5図はDRAMにおける高速アクセス機能を示した波形図、第6図は従来のキャッシュ機能を有するメモリシステムにおけるDRAM素子の構成説明図、第7図は従来のキャッシュ機能を有するメモリシステムのブロック構成図、第8図は従来のキャッシュ動作を示す波形図である。

図において、5はメモリセルアレイ、6はセンスアンプ、22a～22pはラッチ、24はコンパレータ、31a～31dはトランスファゲート、32a～32dはスタティックメモリセルアレイ、34a～34dはブロックデコーダ、35はウェイデコーダ、37はウェイロックである。

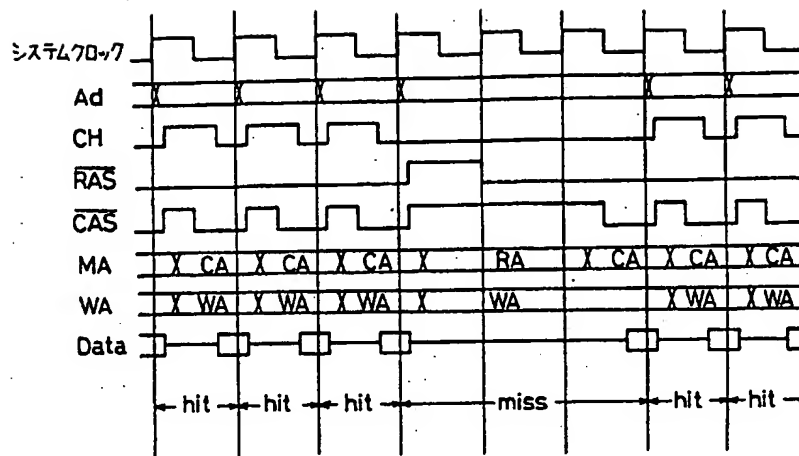
なお、各図中同一符号は同一または相当部分を

示す。

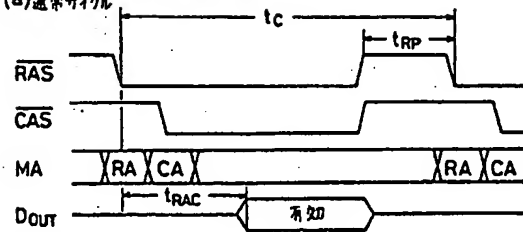
代理人 大 岩 増 造



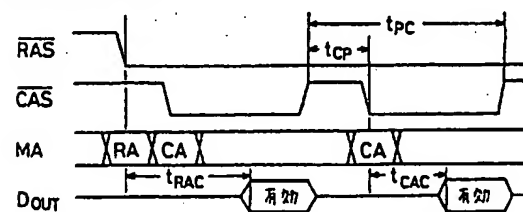
第 4 図



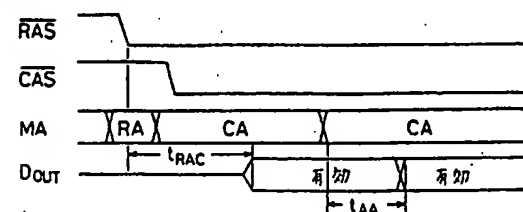
第 5 図
(a)通常サイクル



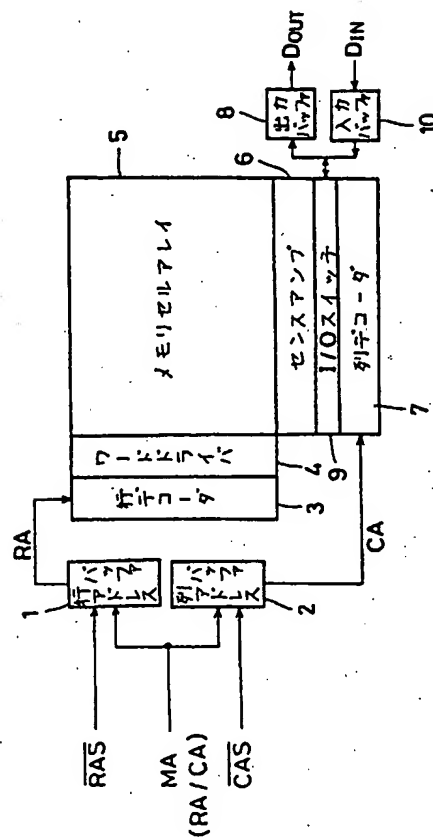
(b)ページモードサイクル



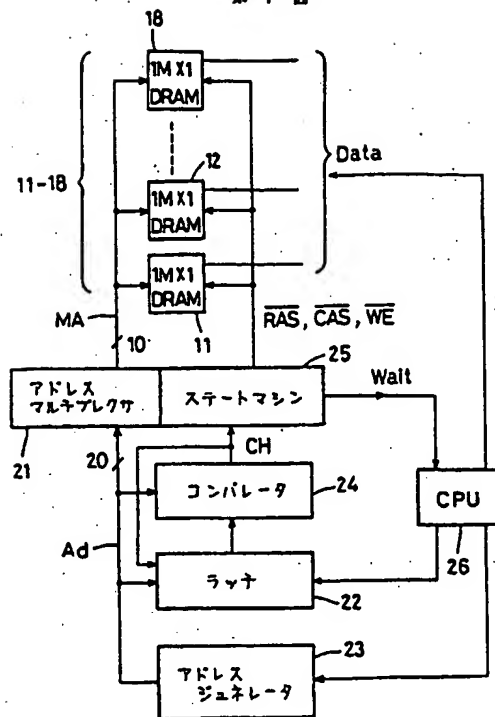
(c)スタティックコラムモードサイクル



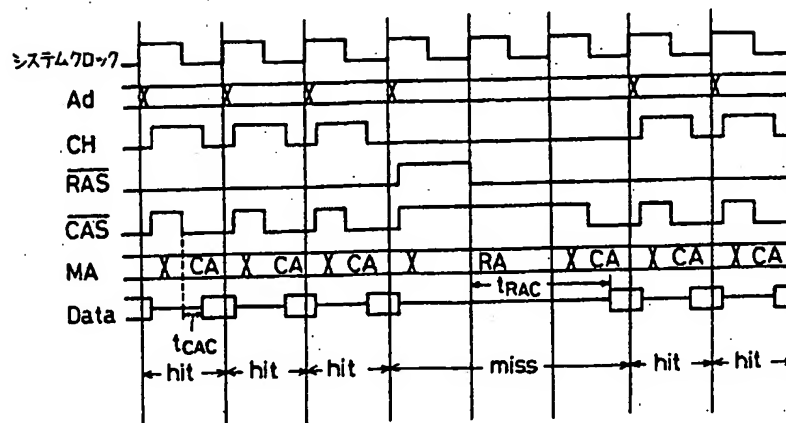
第 6 図



第 7 図



第 8 図



(19) JAPANESE PATENT OFFICE (JP)

(12) PUBLICATION OF UNEXAMINED (KOKAI) PATENT APPLICATION (A)

(11) Kokai (Laid-Open) Patent Application Number 1-159891

(43) Date of Disclosure: June 22, 1989

(51) Int. Cl. ⁴	Identification Symbol	Intra-Agency Number
G 11 C 11/34	362	C-8522-5B

Examination requested: not yet requested
Number of inventions: 1 (total of 10 pages)

(54) Title of the Invention: SEMICONDUCTOR MEMORY DEVICE

(21) Application Number: 62-322126

(22) Filing Date: December 17, 1987

(72) Inventor: Kazuyasu FUJISHIMA
c/o Mitsubishi Electric Corporation
LSI Research Institute
Hyogo-ken, Itami-shi, Mizuhara 4-chome, 1-banchi

(72) Inventor: Yoshio MATSUDA
c/o Mitsubishi Electric Corporation
LSI Research Institute
Hyogo-ken, Itami-shi, Mizuhara 4-chome, 1-banchi

(72) Inventor: Mikio ASAKURA
c/o Mitsubishi Electric Corporation
LSI Research Institute
Hyogo-ken, Itami-shi, Mizuhara 4-chome, 1-banchi

(71) Applicant: Mitsubishi Electric Corporation
Tokyo-to, Chiyoda-ku, Marunouchi 2-chome, 2-ban, 3-go

(74) Representative: Masuo OIWA, patent attorney, 2 others

Specifications

1. Title of the Invention: Semiconductor Memory Device
2. Scope of the Patent=s Claims

(1) A semiconductor manufacturing device, characterized by the fact a semiconductor device, having a memory cell array comprising memory cells in which various types of information are stored in rows and columns, so that information is acquired from said memory cells by specifying row addresses and column addresses;

is equipped with a sense amplifier, which detects and stores information in 1 row segments of a memory cell with a specified row address,

with a static memory cell array, having a plurality of respective memory cells, provided per 1 bit of information contained in this sense amplifiers, which acquires selectively information contained in said sense amplifier,

with a switch means, which has access to said static memory cell array or any of said memory cell arrays;

and with a way decoder, which selects and has access to any of the memory cells in said static memory cell arrays.

(2) The semiconductor memory device of claim 1, characterized by the fact that information contained in said sense amplifier of said static memory cell array is acquired in predetermined block units, further equipped with a block decoder, which has access to any of the blocks in said static memory cell array.

(3) The semiconductor memory device of claim 1 or 2, characterized by the fact that said way decoder is equipped with a function applying the non-activating state to all memory cells in said static memory cell array.

3. Detailed Explanation of the Invention

(Sphere of Industrial Use)

This invention relates to a semiconductor memory device which is provided with a built-in cache memory.

(Prior Art Technology)

In order to increase the cost-performance ratio of conventional computer system, a low-cost and low-speed dynamic RAM (DRAM) was often used in the past for a high-capacity main memory, while a low-capacity and high-speed memory was deployed with a high-speed CPU buffer for the main memory

The high-speed buffer mentioned above, called a cache memory, holds a copy of a block of data requested by the CPU from the main memory. The address data which is accessed by the CPU is acquired, when it is present in the cache memory (a cache hit), from the cache memory which contains the data requested by the CPU. On the other hand, when the data being accessed by the CPU is not present in the cache memory (a cache miss), the CPU will acquire the requested data from the low-speed main memory (DRAM).

Because expensive high-speed memory is needed to incorporate the cache memory system described above in the main memory system, it has not been possible to use such memory in compact computer systems due to the cost, as expensive high-speed memory was required. Therefore, a simple cache system construction was created utilizing a high-speed access function of DRAM having the page function, or the static column mode function, etc.

The reference below in the waveform diagram shown in Figure 5 explains the performance of the page mode and the static column mode operations. As shown in Figure 5, diagram (a) indicates a normal DRAM cycle, (b) indicates the page mode, and (c) indicates the static column mode cycle.

As shown in Figure 5 (a), during the normal cycle, at the falling edge of the RAS (Row Address Strobe) signal, the row address RA (Row Address) is incorporated in the DRAM from the multiplex address signal MA, while the column address CA (Column Address) is incorporated in the DRAM from the multiplex address signal MA at the falling edge of the CAS (Column Address Strobe) signal. After that, the data output D_{OUT} of selected memory cells is obtained from the row address RA and column address CA. Because the data is read in the above-described cycle during the normal cycle, a time period is required from the point of the falling edge of the RAS signal to the point in time t_{RAC} (RAS access time) when the data output D_{OUT} becomes valid. This access time period t_{RAC} is normally about 100 ns. In addition, t_{RP} is the precharge time period of signal of the RAS signal and t_C is the time period of the cycle, which normally equals about $t_C = 200$ ns.

Diagram (b) indicates the situation when data is read in a plurality of column addresses CA on the same row on address RA in the page mode cycle. Accordingly, a time period t_{CAC} (CAS access time) is created from the point of the falling edge of the CAS signal to the point when the data output D_{OUT} becomes valid as a time corresponding to about half of the access time period t_{RAC} during a normal cycle. In addition, t_{CP} is the precharge time period of the CAS signal, and t_{PC} is the cycle time period.

Diagram (c) indicates the situation when the falling edge of the page mode CAS signal is not required in static column mode, since the operation is conducted as with column address CA in static RAM. Accordingly, an access time period t_{AA} (address access time) is created from the point when the multiple address MA is changed to the point when the data output D_{OUT} becomes valid, which corresponds to about a half of the access time period t_{RAC} during a normal cycle of the same t_{CAC} , which is normally about 50 ns.

Figure 6 is a construction block diagram which shows the basic construction of a DRAM

element according to prior art enabling page mode or static column mode.

As shown in the same figure, the construction includes a row address buffer 1 and a column address buffer 2, incorporating multiplex address signal RA and column address CA. In addition, when the falling edge of the RAS signal is input to the row address buffer 1, the row address signal is sent to a row decoder 3. At the next stage, when a word driver 4 is operated, 1 word line (not shown in the figure) contained in a memory cell array 5 will be activated by the row address signal RA.

Also, the data of the memory cells connected to an activated word line is sent through all the bit lines (not shown in the figure) from the memory cell array 5 to a sense amplifier 6. The sense amplifier 6 detects and amplifies obtained data. Accordingly, the data in one line of the line address RA specified at this point in time is latched by the sense amplifier 6. After that, when the same data is accessed in the line address RA, the page mode and the static column mode described above can be utilized.

[page 3]

Specifically, during the page mode, when the falling edge of the CAS signal is input to the column address buffer 2, the column address CA is sent to a column address decoder 7 and because one of the data groups stored in the sense amplifier 6 becomes valid, data output D_{OUT} will be obtained through the output buffer 8. The operations are performed in the same manner also during the static column mode, except for the point when the multiplex address MA is changed. In addition, 9 is an I/O switch controlling the input and output of data, and D_{IN} is a data input.

Figure 7 is a block construction diagram showing a conventional memory system having a simple cache system utilizing the page mode (or static column mode). As shown in the same figure, this memory system is a 1 M byte memory system constructed so that 8 DRAM elements 11 ~ 18 are used with 1 M x 1 construction of 8 individual elements. Accordingly, while 20 address lines are required ($2^{20} = 1948576 = 1 \text{ M}$), multiplex signal MA is in reality sent from an address multiplexer 21 divided to column address CA (10 bits) by 10 address line to respective connected DRAM elements 11 ~ 18.

Figure 8 is a waveform diagram indicating the cache operation of the memory system shown in Figure 7. The following is an explanation of the operation of the memory system shown in Figure 7 with reference to Figure 8 and Figure 6. Moreover, it is assumed that row address RA1 is latched by latch 22 just prior to being accessed, and that all the data in the row address RA1 contained in the sense amplifier 6 has been latched.

During this state, address signal A_d containing 20-bit data is generated by the generator 23 as required data. When the row address data RA2 obtained from this address signal A_d is input to converter 24, this row address RA2 is compared to the row address RA1, which is stored

in the latch 22, by the comparator 24, and if $RA1 = RA2$, the data group held in the sense amplifier 6 will be accessed (a cache hit).

The comparator 24 will therefore send activated ("H" level) cache hit signal CH (Cache Hit) to a state machine 25. When the state machine 25 receives an activated CH signal, the RAS signal is maintained on the "H" level, the page mode is controlled by toggling of the CAS signal (rising is followed by falling), the address multiplexer 21 supplies the column address CA as multiplex address MA to the DRAM elements 11 ~ 18, and selected data is acquired from the column decoder 7 from the data groups stored by the sense amplifier in each of the DRAM elements 11 ~ 18. When a cache hit occurs in this manner, output data D_{out} can be obtained with the high-speed access time t_{CAC} from the DRAM elements 11 ~ 18.

On the other hand, when it is determined by the comparator 24 that $RA1 \neq RA2$, because a data group has been accessed which is outside of the data group that is held in the sense amplifier 6, the comparators 24 will generate signal deactivating the CH signal ("L" level) for the state machine 25. At this time, the state machine 25 will exercise control over the DRAM elements 11 ~ 18 so that the normal cycle will be applied sequentially to the RAS and CAS signal, and the address multiplexer 21 will supply the multiplex address MA with the sequence of the column address CA to the DRAM elements 11 ~ 18. When a cache miss has occurred in this manner, precharging is carried out with the RAS signal as shown in Figure 8, and output data D_{out} will be obtained with the low-speed access time period t_{RAC} from the DRAM elements 11 ~ 18. Because of that, the state machine 25 will generate the wait signal Wait and the standby mode is applied to the CPU 26. In addition, when the latch 22 receives cache hit signal CH which has not been activated by the comparator 24, a new row address 22 will be held.

(Problems to Be Solved by the Invention)

With the conventional simple cache system that was provided with a construction of the type in which latching was applied with the sense amplifier described above, 1 entry was available.

[page 4]

Accordingly, because a cache hit was enabled only if there was a connection to the same row address RA, a cache miss was in the end thus generated for example when there were 2 row addresses or when repeated operations were carried out with an extended program routine, creating a problem known as a low cache hit rate.

The purpose of this invention is to solve the problem described above by providing a semiconductor memory device which has a simple cache hit system enabling to increase the cache hit system.

(Means to Solve Problems)

The semiconductor memory device of this invention provides a construction, applied to a memory cell array comprising memory cells in which information is stored in arrays in the shape of rows and columns, with a method wherein that the information is acquired from said memory cells when the row address and column address is specified;

equipped with a sense amplifier, which detects and stores information in 1 row segments for a specified row address;

with a static memory cell array, having a plurality of memory cells, provided for each 1 bit of information contained in this sense amplifier, and which selectively incorporates in the memory cells respective information items contained in said sense amplifier;

with a switch means, which enables in a selective manner access to one of said memory cell arrays;

and with a way decoder, which selects whether and which of the memory cells is to be accessed in said static memory cell array.

(Operation)

Because the static memory cell array is provided with a plurality of memory cells per 1 bit of information contained in a sense amplifier, the data can be held in different row addresses.

(Embodiment)

Figure 1 is a block diagram showing the basic construction of the DRAM elements of a memory system having a cache function in one embodiment of this invention. Because as shown in this figure, items 1 ~ 4, 8 ~ 10, and CAS, RAS, MA, RA, CA, CH are identical to the prior art example, an explanation of these items are omitted and only points which are different from the conventional example are explained below.

As shown in the same figure, a memory cell array 5 is used, which is divided into 4 block segments B1 ~ B4, wherein a transfer gate 31 corresponding to the blocks B1 ~ B4, and a static cell memory array 32 (32a ~ 32d) are inserted between a sense amplifier 6 and an I/O switch 9. Because the transfer gate 31 is controlled with a block decoder 34 using the block configuration shown in detail in Figure 2, transmission of data in the memory cell array 5 to the static memory cell arrays 32a ~ 32d is enabled in block units (B1 ~ B4) through the corresponding part of the sense amplifier 6 when a conductive or non-conductive status of this configuration is created.

As shown in Figure 2, in the static memory cell array 32 are deployed 4 static memory cells 32_{m1} ~ 32_{m4} via the transfer gate 31, corresponding to 1-bit information stored in the sense amplifier 6. These static memory cells 32_{m1} ~ 32_{m4} are activated by the "H" level of the output lines W₁ ~ W₄ of the way decoder 35.

As shown in Figure 1, the way decoder 35 decodes the way an address WA is input through the way address buffer 36, so that the "H" level is selectively applied to the output lines $W_1 \sim W_4$.

The block decoders 34a ~ 34d are controlled by activation applied by an AND gate G1 when input signal is created with inverted signal to the signal CH and to the upper 2 bits of the column address CA. Specifically, when the CH signal is at the "L" level, one of the block decoders 34a ~ 34d, selected by the upper 2 bits of the column address CA, is activated. When the CH signal is at the "H" level, none of the block decoders 34a ~ 34d is activated. In addition, when one of the block decoders 34a ~ 34d is activated, the corresponding transfer gate 31a ~ 31d is rendered conductive. On the other hand, when the column decoder 7 inputs the column address signal CA, one of the I/O switches 9 is rendered valid.

Figure 3 is a block diagram showing the construction of a memory system having the cache function in one embodiment of this invention. As shown in the same figure, unlike according to prior art, a construction is created containing 4 blocks, with 4 static memory cells 32m1 ~ 32m2, 16 latches 22a ~ 22p are deployed (22a ~ 22d for block B1, 22e ~ 22h for block B2, 22i ~ 22l for block B3, and 22m ~ 22p for block B4).

[page 5]

Also, these latches 22a ~ 22p are used to store way addresses WA, indicating which item is to be selected in respective row addresses RA and static memory cells 32m1 ~ 32m4 so that 4 items can be selected in block units from the upper 2 bits of the column address CA of the address signal A_4 .

The comparator 24 incorporates the row addresses and compares the to 4 stored row addresses selected from the latches 22a ~ 22p, so that if there is a coincidence, this is considered as a cache hit in 1 latch and the cache hit signal CH is activated (at the "H" level) and output. At the same time, the way address C_{WA} , stored in the coinciding latch 22, is output to the way logic 37. On the other, if there is no coincidence in any of the 4 selected latches, this is considered a cache miss and the deactivated cache hit signal CH (at the "L" level) will be output.

When the cache hit signal CH and way address signal C_{WA} are input as an input signal from the comparator 24, when there is a cache hit, the way logic 37 outputs the input way address C_{WA} as is in the form of the way address WA to each DRAM 11 ~ 18. On the other, when there is a cache miss, the way address WA specified according to a specific algorithm is output to each DRAM 11 ~ 18 and latch 22a ~ 22p. However, other methods are also conceivable, for example the simple first input - first output method (FIFO: First-In, First-Out), or the method based on the access time interval of the oldest item (LRU: Least Recently Used), or a similar method.

The cache hit and cache miss operation of the memory system in one embodiment of this

invention as shown in Figures 1 ~ 3 will now be explained with reference to the waveform diagram shown in Figure 4. In addition, it is assumed that just before the latches 22a ~ 22p have been accessed, the corresponding wait address of row addresses RA1a ~ RA1p and respective latches 22a ~ 22p were latched, and that all of the data in each memory cell 32_{m1} ~ 32_{m4} of the static memory cell arrays 32a ~ 32p was already latched at that time in each block B1 ~ B4.

In a similar state, when this is requested by the CPU, not shown in the figures, address signal Ad having 20 bits will be generated by the address generator 23. The row address RA2 will be output from this address signal A_d to the converter 24. On the other hand, only 4 of the latches 22a ~ 22p, corresponding to the blocks B1 ~ B4 selected by the upper 2 bits in the column address CA of the address signal A_d, will be valid. In this case, it will be assumed for explanation purposes that when the latches 22e ~ 22h of the block B2 have been selected, a comparison is performed of the input row address RA2 and of each row address RA1e ~ RA1h stored in the latches 22e ~ 22h by the comparator 24. If one of the addresses RA1e ~ RA1h coincides with RA2, this will be considered a cache hit, the cache hit signal CH will be activated (at the "H" level) and sent to the state machine 25, the way logic 37 and each DRAM element 11 ~ 18. If the row address coincided with RA2, the way address signal C_{WA}, which is stored in the latches 22e ~ 22h, is sent to the way logic 37. When the cache hit signal CH is received at the "H" level by the way logic 37, the input way address C_{WA} will be output as is in the form of the way address WA to each DRAM 11 ~ 18.

Because the "H" level of the CH signal has been created at this time, the block decoder 34 will not be activated at all, the transfer gate 31 will be rendered completely non-conductive, and the electrical connection between the static memory array 32 and the sense amplifier 6 will be interrupted.

On the other hand, the state machine 25 performs page mode control by toggling of the CAS signal, and the address multiplexer 21 furnishes the column address CA as the multiplex address MA to the DRAM elements 11 ~ 18.

[page 6]

Moreover, the way address WA input to each DRAM element 11 ~ 18 will be input through the way address buffer 36 to the way decoder 35. Also, when the way address WA is decoded by the way decoder 35, the level of the output line W_i is raised to the "H" level (wherein i equals one of the items 1 ~ 4). When this occurs, the data of the column address selected by the column decoder 7 from the memory cell 32_{mi}, connected to the output line W_i in which the level was raised to the "H" level inside the static memory cell array 32b of each DRAM element 11 ~ 18, is output through the I/O switch 9. When a cache hit has occurred in this manner, output data D_{out} can be obtained with the high-speed access time period T_{CAC}.

Also, when no coincidence has been determined by the comparator 24 in RA2 and RA1e ~ RA1h, since this is considered a cache miss, non-activating cache hit signal CH (at the "L"

level) will be sent to the state machine 25, wait logic 36 and each DRAM element 11 ~ 18. When the cache hit signal CH has been received at the "L" level, the way logic 37 outputs the way address specified according to a specific algorithm to each DRAM element 11 ~ 18 and to the latches 22a ~ 22p.

Because the CH signal is created at the "L" level at this time, only the block decoder 34b will be activated, the transfer gate 31b is rendered conductive, and an electrical connection is created between the static memory cell array 32b and sense amplifier 6. In addition, the electrical connection is disconnected between the other static memory cell arrays 32a, 32c, 32d and the sense amplifier 6.

On the other hand, when the RAS signal is falling, next, the state machine 25 applies control to the DRAM element 11 ~ 18 in the falling cycle of the CAS signal, and the address multiplexer 21 supplies the multiplex address MA in the sequence row address RA2, column address CA, to the DRAM elements 11 ~ 18. Also, the way address WA is input through the way address buffer 36 to the way decoder 35 of each DRAM element 11 ~ 18. Further, because the way address WA is decoded by the way decoder 35, the level of the output line W_i (wherein i equals one of the items 1 ~ 4) is raised to the "H" level.

Also, the data selected by the column decoder 7 through the sense amplifier 6, transfer gate 31b and static memory cell array 32b, I/O switch 9 and output buffer 8 from the memory cell array 5, is read as output data D_{out} . At the same time, data contained in the memory cell 32_{mi} , which is connected to the output line W_i in which the level was raised to the "H" level inside the static memory cell array 32, is rewritten. During a similar cache miss, output data D_{out} is therefore obtained with the low-speed access time period T_{RAC} from the DRAM elements 11 ~ 18. Because of that, the state machine 25 generates the wait signal Wait and the wait state is applied to the CPU 26. Further, a new row address RA2 is held in one of the latches 22e ~ 22h selected by the way address WA and by the upper 2 bits of the upper column address CA (the value of the other latches 22 remains unchanged).

In this manner, because 1-bit information contained in the sense amplifier 6 holding information corresponding to 1 row segment can be stored in any of the 4 static memory cells $32_{m1} \sim 32_{m4}$, 4 entries are available in the static memory cell array 32. As a result, the cache hit rate is increased in cases such as when the program routine is realized repeatedly so that it is continuously extended to 2 row addresses, etc.

Further, because the memory logic of the DRAM elements 11 ~ 18 can be realized in the block units B1 ~ B4 during a cache hit or during a cache miss, data groups corresponding to row addresses can be stored independently per each block B1 ~ B4 in the static memory cell arrays 32a ~ 32d, which means that the number of entries corresponding to $4 \times 4 = 16$ is created, enabling to increase the cache rate even more.

Furthermore, the invention can be applied not only to memory reading and writing operations as indicated in the embodiment shown in Figure 1. For example, when write signal WE is input to the way decoder 35 so that all the output lines $W_1 \sim W_4$ are set to the

non-activating state during writing operations ($\overline{WE} = "L"$), all the memory cells $32_{m1} \sim 32_{m4}$ in the static memory cell array can be set to the inactive state regardless of the value of the way address WA, and regardless of whether the CH signal is at the "H" level or the "L" level.

[page 7]

Further still, although an embodiment was explained in which the storage of 1-byte information of the sense amplifier was enabled to any of 4 static memory cells $32_{m1} \sim 32_{m4}$, this number can be suitable increased or decreased. Also, while a construction comprising 4 blocks B1 ~ B4 was used for the memory cell array 5 in this embodiment, it goes without saying that the number of the block segments can be also increased or decreased as appropriate.

(Effect of the Embodiment)

As was explained above, because multiple memory cells are deployed for 1-byte information contained in a sense amplifier with the static memory cell array of this invention, data can be held on different row addresses. The result is that the cache rate can be increased because the number of entries is increased.

4. Brief Explanation of Figures

Figure 1 is a diagram explaining the construction of DRAM memory elements in one embodiment of this invention, Figure 2 is a diagram explaining in detail the construction of the DRAM element shown in Figure 1, Figure 3 is a block diagram showing the construction of a memory system having the cache function in one embodiment of this invention, Figure 4 is a waveform diagram explaining the cache operation in one embodiment of this invention, Figure 5 is a waveform diagram indicating the high-speed access function in a DRAM device, Figure 6 is a diagram explaining the construction of a DRAM element in a memory system having the cache function according to prior art, Figure 7 is a diagram showing the block construction of a memory system having the cache function according to prior art, and Figure 8 is a waveform diagram showing the cache function according to prior art.

In these figures, 5 is a memory cell array, 6 is a sense amplifier, 22a ~ 22p are latches, 24 is a comparator, 31a ~ 31d are transfer gates, 32a ~ 32d are static memory cell arrays, 34a ~ 34d are block decoders, 35 is a way decoder, and 37 indicates a way logic.

Also, the same symbols are applied to the same or corresponding parts.

Representative: Masuo OIWA, patent attorney.

Figure 3

- 21 address multiplexer
- 23 address generator
- 24 comparator
- 25 state machine
- 37 way logic

[page 8]

Figure 1

- 1 row address buffer
- 2 column address buffer
- 3 row decoders
- 4 word driver
- 5 memory arrays B1, B2, B3, B4
- 6 sense amplifier
- 7 block decoder
- 8 output buffer
- 10 input buffer
- 31 transfer gates
- 32 memory cell array
- 34a block decoder
- 34b column decoder
- 35 way decoder
- 36 way address buffer

- 32 static memory cell array

Figure 2

- (A) block decoder
- (B) I/O switch
- (C) static memory cell
- (D) transfer gate
- (E) block boundary

[page 9]

Figure 4

(A) system clock

Figure 5

- (a) normal cycle
- (b) page mode cycle
- (c) static column mode cycle
- (1) valid
- (2) valid
- (3) valid
- (4) valid

Figure 6

- 1 row address buffer
- 2 column address buffer
- 3 row decoder
- 4 word driver
- 5 memory cell array
- 6 sense amplifier
- 7 column decoder
- 8 output buffer
- 9 I/O switch
- 10 input buffer

[page 10]

Figure 7

- 21 address multiplexer
- 22 latch
- 23 address generator
- 24 comparator
- 25 state machine

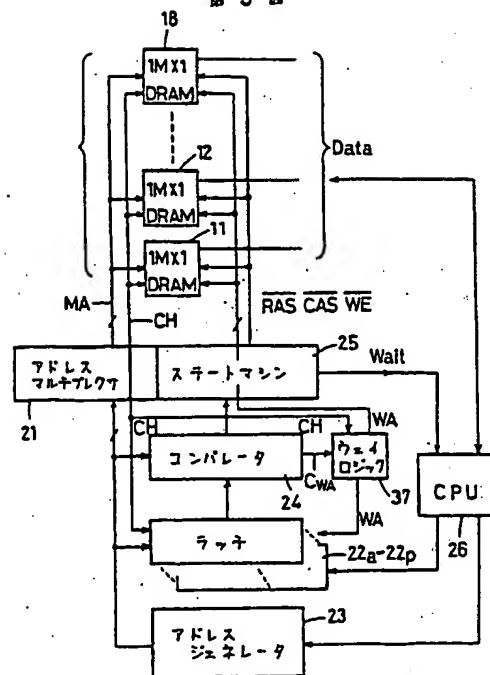
Figure 8

(A) System clock

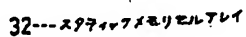
示す。

代理人 大 岩 雄 雄

第 3 図



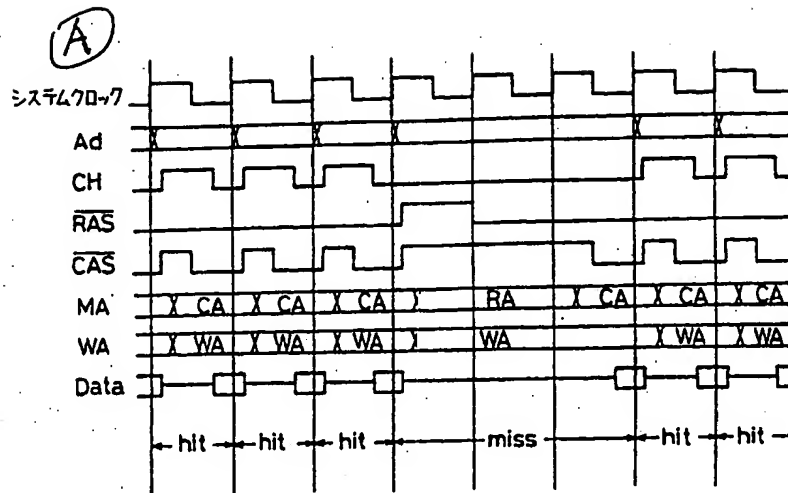
8



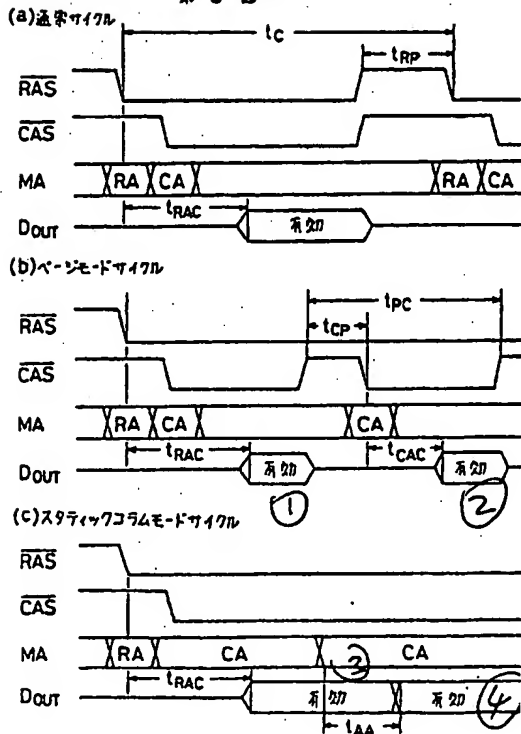
第 2 圖



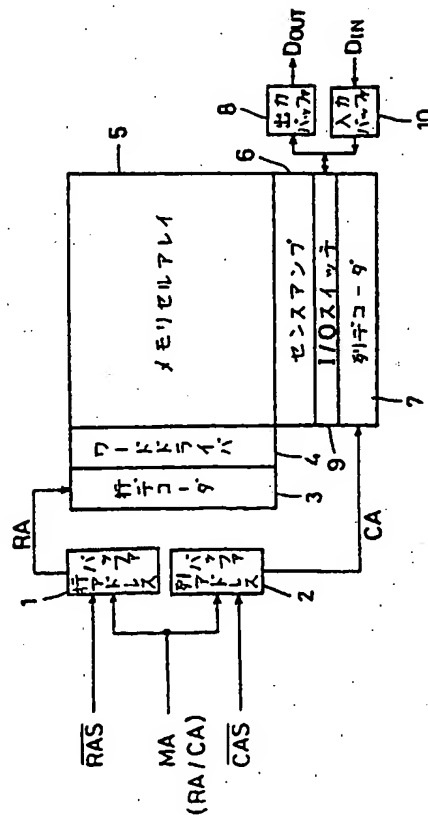
第 4 図



第 5 図

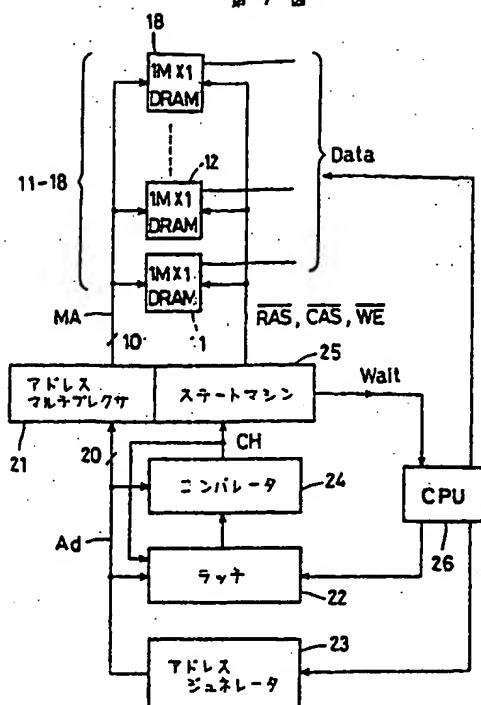


第 6 図



10

第 7 図



第 8 図

